# Improving Stormwater Program Monitoring, Tracking, Evaluation, and Reporting

## Workshop Report and Recommendations for Program Improvement

Report Date: June 29, 2018

****INITIAL DRAFT****

*Prepared for:*

U.S. Environmental Protection Agency
Region 9

75 Hawthorne Street
San Francisco, CA 94105

*Prepared by:*

PG Environmental

14555 Avion Parkway, Suite 125
Chantilly, VA 20151

*EPA Contract No. EP-C-16-003*

## ACKNOWLEDGEMENTS

EPA Region 9 appreciates its partnership with the State of California and thanks the State for hosting and co-sponsoring the workshop. This workshop was intended to advance efforts by EPA and the State to improve permitting, promote water technology innovation, and help enhance the effectiveness and long-term viability of urban water management programs. It would not have been possible without the financial support of EPA Headquarters' Office of Science and Technology and Office of Wastewater Management. EPA also acknowledges PG Environmental's work in supporting the workshop and developing this report.

Special thanks must also be extended to the workshop participants from across the country who energetically and thoughtfully engaged in the workshop and in the preparation of this report.

## DISCLAIMER

The material presented in this document is intended solely for informational purposes and is not intended to interpret federal statutory or regulatory requirements or to convey EPA guidance or policy. The recommendations and possible actions described in this document do not represent binding commitments by EPA or other parties. This document is not intended, nor can it be relied on, to create any rights enforceable by any party in litigation with the United States. This document may be revised or updated without public notice.

Mention of trade names or commercial products in this document does not constitute an endorsement or recommendation for use.

# TABLE OF CONTENTS

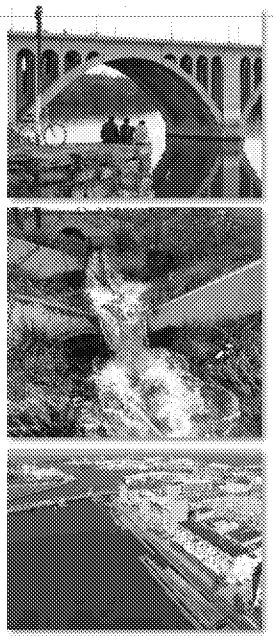
# EXECUTIVE SUMMARY

Commented [AK1]: Note to EPA: We will build out the Executive Summary after the report content is stable.

Commented [AK2]: Note to EPA: The images in this report are placeholders; they will be updated in a subsequent version.

This report aims to provide a synthesis of participant ideas and contributions along with other existing research to identify the most impactful opportunities for strengthening Municipal Separate Storm Sewer System (MS4) permits and program implementation. It includes an overview of the workshop discussions, specific actions identified, case studies, summaries of known efforts related to the recommendations, and some indication of commitment by groups to make progress related to a given recommendation. The table on the following pages presents a brief synopsis of these recommendations.

The U.S. Environmental Protection Agency (EPA), the State of California, and participating organizations plan to build upon workshop conversations through broad outreach to partners and stakeholders and continued dialogues surrounding these important issues. This iterative, inclusive approach allows for objective evaluation of program progress to date, assessment of opportunities for program adjustment to better meet Clean Water Act (CWA) goals, and identification of specific actions necessary to enable new, innovative permitting approaches across the nation.



Photos (top to bottom): EPA, PG Environmental, stock

| Recommendations for Capacity Building and Program Support |
|---|
| • Develop a Vision for the Future of Stormwater Monitoring to Improve Program Efficiency and Effectiveness.<br>• Develop Guide to Improving Monitoring and Evaluation to Better Serve MS4 Programs.<br>• Establish Key Performance Metrics (Activity- and Outcome-Based) for Municipal Stormwater Programs.<br>• Leverage Existing Data Sets to Enhance Approaches for Informing Program Management Decisions and Gauging Program Effectiveness. |
| Permitting Recommendations |
| • Improve Clarity of Monitoring and Effectiveness Permit Requirements and Objectives and Methods/Designs.<br>• Create a Pathway in Permits to Make Special Studies More Impactful.<br>• Identify Instances When Lack of Approved Monitoring Methods Created a Barrier to Technology Implementation. |
| Making Outfall and Receiving Water Monitoring More Discriminating to Inform Program Management |
| • Evaluate Options and Effectiveness of Conducting Monitoring Efforts at Scales Designed to Yield Actionable Results.<br>• Convene Visioning Session for Deploying Sensors in Municipal Stormwater Programs. |
| Improving Our Ability to Quantify Effectiveness—Approaches to Link Water Quality Outcomes to Actions |
| • Document Current State of Knowledge of BMP Performance and Effectiveness.<br>• Improve the Applicability and Usefulness of Modeling through Collecting and Incorporating Better BMP Performance Data.<br>• Establish Model Calibration Guidance to Improve the Usefulness of Models.<br>• Evaluate Methods to Account for True Source Controls in Models. |
| Improving Program Tracking and Reporting |
| • Identify an Approach for Using Established Performance Metrics to Guide Tracking and Reporting Efforts.<br>• Determine the Most Effective MS4 Program Reporting Mechanisms and Formats. |

**Commented [AK3]:** Note to EPA: This table will be expanded when the body text is more stable.

# Abbreviations and Acronyms

Commented [BJ4]: Note to EPA: This will be fully updated during editorial review once the body text is stable after revisions.

| | |
|---|---|
| BMP | best management practice |
| CASQA | California Stormwater Quality Association |
| CFR | Code of Federal Regulations |
| CWA | Clean Water Act |
| EPA | U.S. Environmental Protection Agency |
| MCM | minimum control measure |
| MS4 | Municipal Separate Storm Sewer System |
| NMSA | National Municipal Stormwater Alliance |
| NPDES | National Pollution Discharge Elimination System |
| POC | pollutants of concern |
| RAA | reasonable assurance analysis |
| SWMM | Storm Water Management Model |
| SWMP | stormwater management program |
| TMDL | total maximum daily load |
| WEF | Water Environment Federation |
| WQS | water quality standards |

# 1 INTRODUCTION


Photo: EPA

The U.S. Environmental Protection Agency (EPA) established the Municipal Separate Storm Sewer System (MS4) program in the early 1990s under its National Pollution Discharge Elimination System (NPDES) permits to address pollution from urban stormwater runoff. Now almost 30 years later, regulators and permittees have a greater understanding of urban water quality management, have compiled illustrative examples of program successes and failures, and are using new technologies for data management, modeling, and water quality monitoring.

While the MS4 program has evolved in response to new information and tools, significant opportunities for improvement remain. EPA convened a small group of stakeholders in late 2017 to assess the MS4 program at large and identify the most impactful opportunities for strengthening permits and building program capacity. This report presents the discussions and ideas from a follow-on workshop in March 2018 that focused on approaches to monitoring, tracking, evaluation, and reporting.

In December 2017, EPA convened a workshop that focused on MS4 program minimum control measures, industrial program requirements, and water quality-based control requirements. The resultant white paper, ***Evolution of Stormwater Permitting and Program Implementation Approaches***, captures workshop discussion and recommendations for program improvement and provides background information regarding the overall MS4 program.

In recent years, permittees have expressed the need to reassess these approaches to improve MS4 program implementation. For example, many programs across the country have found it difficult to establish monitoring and evaluation efforts that illustrate the direct impact of program actions on water quality outcomes. Further, there is wide variability with the monitoring and evaluation requirements included in MS4 permits and the relative amount of resources expended on each. As a result, much time and money is spent on monitoring efforts that are not designed to answer key questions regarding program effectiveness. With no widely-accepted approach to assessing effectiveness, workshop participants concurred that there is an opportunity to create a better mix of water quality monitoring, evaluation, tracking and reporting programs.

> "Monitoring should be a way to change incrementally the standard — not punish the willing. Management and permitting actions must evolve as experience leads to opportunities for improved practice and better-informed expectations" (Water Environment Federation [WEF], 2015, pg. 22).

This report provides a synthesis of workshop participant ideas for improvements to monitoring, tracking, evaluation, and reporting along with other existing research. The full set of recommendations, presented in Section 4, includes discussion overview, related actions, case studies, and some indication of commitment by groups to make progress related. Inclusion of a recommendation does not necessarily indicate the support of all participants; rather, it provides an opportunity for further discussion, inquiry, and possible progress.

# 2 MS4 WORKSHOP

In March 2018, EPA Region 9, with assistance from EPA Headquarters and in partnership with the State of California, invited 31 stormwater experts from across the country to Oakland, California, for a two-day workshop titled *Improving Stormwater Permit Approaches to Monitoring, tracking, evaluation, and reporting* (full list of workshop participants included in Appendix A). The workshop was designed to explore current requirements and practices for municipal stormwater program monitoring, tracking, evaluation, and reporting and identify opportunities for improvement that would support more effective program implementation. Importantly, primary goals were to identify (1) how permits can direct or incentivize these improvements, (2) what methods could be used to support these improvements (e.g., guidance, best practices, research), and (3) what entities within the sector could help affect these changes.

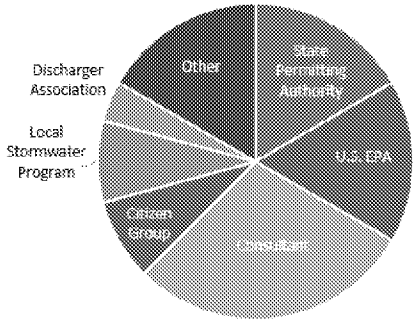



**Figure 1.** Relative distribution of workshop participants across the sector.

Through facilitated dialogues, participants were asked to reflect on their own first-hand experiences with MS4 permitting and program implementation as it relates to monitoring, tracking, evaluation, and reporting. To promote honesty and openness, participants agreed that the viewpoints expressed would not be attributed to individuals in this resultant report.

As noted above, this workshop was a follow-on to a prior event which focused on MS4 program minimum control measures, industrial program requirements, and water quality-based control requirements. The resultant white paper, *Evolution of Stormwater Permitting and Program Implementation Approaches*, captures workshop discussion and recommendations for program improvement, and provides background information regarding the overall MS4 program. Thus, this report does not duplicate the background information and focuses more directly on the workshop discussions and recommendations for improvement to monitoring, tracking, evaluation, and reporting.

## 2.1 Workshop Format

Through a facilitated dialogue, invited representatives from federal, state, and local government, as well as sector stakeholders (e.g., permit holders, trade associations, nonprofit organizations), evaluated MS4 program monitoring, tracking, evaluation, and reporting approaches to inform possible changes in NPDES permit provisions and opportunities to improve MS4 programs.

The workshop included 7 sessions over two days in a format designed to efficiently identify recommendations specific to monitoring, tracking, evaluation, and reporting approaches (full agenda included in Appendix B).

**Workshop Sessions**

1. Current Condition - Are the current Monitoring, Evaluation, Tracking and Reporting requirements effective?
2. How Can We Better Use Performance Metrics to Facilitate Improved Monitoring, tracking, evaluation, and reporting?
3. How Can We Make Outfall and Receiving Water Monitoring More Useful?
4. Linking Activities to Expected Water Quality Outcomes
5. How Can We Improve Program Performance Tracking?
6. Reforming Reporting Approaches to Help Move Programs Forward and Give Permitting Authorities What They Need
7. Reflection, Synthesis, and Wrap Up

Each workshop session followed the same general structure with a conversation starter, facilitated group discussion, and identification of important findings and specific actions discussed to strengthen and improve approaches to monitoring, tracking, evaluation, and reporting.

This report captures the essence of these conversations so that others may benefit from the collective expertise. EPA plans to continue working with various partners and stakeholders to refine and implement the most promising ideas for strengthening MS4 programs through improved monitoring, tracking, evaluation, and reporting, and enabling new, innovative permitting approaches.

## *2.2 Pre-Workshop Questionnaire*

In advance of the workshop, participants were polled to gauge their attitudes toward specific aspects of the permitting program by responding to a series of hypotheses. Twenty-four submissions were received in total. Respondents overwhelmingly agreed that there was potential to realize cost-effective positive environmental outcomes through improved approaches to monitoring, tracking, evaluation, and reporting.[1]

| | Significant or Some Potential | Little or No Potential |
|---|---|---|
| **Water Quality Monitoring** (receiving water, outfall, within collection system, at project or practice scale) | 100 percent | - |
| **Non-Water Quality Evaluation** (activity evaluation, effectiveness evaluation) | 88 percent | 8 percent |
| **Tracking** (tracking discrete activities (e.g., inspections, street sweeping, best management practice [BMP] installation), active asset management planning and tracking) | 100 percent | - |
| **Reporting** (annual reporting to permit authorities, reporting to public or elected officials) | 88 percent | 12 percent |

[1] Participants could also respond "no opinion or insufficient knowledge."

There was strong support for numerous statements on the pre-workshop questionnaire which helped frame onsite discussion and can help orient further consideration of designs for monitoring, tracking, evaluation, and reporting. Select statements are identified below.

| | Strongly Agree or Agree | Strongly Disagree or Disagree |
|---|---|---|
| Monitoring designs must go beyond just data collection methods to include data management, data analysis, and reporting formats that clearly link data collected with Performance Metrics. | 100 percent | - |
| Performance metrics need to be established in concert with improved monitoring designs and methods. | 96 percent | - |
| Metrics should enable evaluation not just of what was done, but also of whether those actions were effective. | 96 percent | - |
| No one monitoring and evaluation method addresses all the assessment needs; multiple approaches tailored to local circumstances are needed. | 92 percent | 4 percent |
| Better guidance and training on new reporting frameworks and how to incorporate them in permits will be needed to advance reporting approaches at the state and local levels. | 92 percent | 4 percent |

Respondents also provided additional insights and suggestions through the pre-workshop questionnaire. One recurring theme was that assessing effectiveness cannot be accomplished through a "one size fits all" approach. Two respondents captured these sentiments as follows:

> *"There isn't one right answer for every program, but there must be a better monitoring/tracking/assessment framework that could be used to build more effective programs across the country."*

> *"Effectiveness assessment is element-specific. No one measurement fits all. So, rather than specifying a measurement, specify a process to follow between the different elements to identify the appropriate measurement, etc. Process would be something like:* ***Inquiry*** *(question, permit requirement, exceedance)* ***→ pollutants of concern (POC) → BMP → Effectiveness measurement → Effectiveness methodology → Report.***"

Additional questionnaire findings are incorporated throughout the report, where applicable; Appendix C summarizes questionnaire results.

## 2.3 Terminology

A purpose of this workshop was to synthesize participant ideas for improvements to monitoring, tracking, evaluation, and reporting along with other existing research. To do this, workshop participants needed to establish a common terminology. It is known that there is variability in terminology across the country that reflects differences in program requirements and approaches; however, the general concepts are relevant to the national program. How MS4 programs (1) track activities, (2) evaluate progress, (3) pose key questions to answer through monitoring, (4) sample stormwater runoff or receiving waters, (5) analyze results, (6) attempt to make program changes in response to observations, and (7) report to permitting authorities—broadly viewed as "monitoring, tracking, evaluation, and reporting." A common set of definitions are presented below for the purposes of this report.

- **Monitoring** – Water quality monitoring typically performed by the permittee at end-of-pipe, in-stream, or in a receiving water.
- **Tracking** – Collecting and compiling information on program implementation (including the use of electronic databases and other systems to document program information) to inform evaluation and reporting.
- **Evaluation** – A determination whether a given program, program element, activity, or individual BMP is meeting its intended objectives.
- **Reporting** – Presenting information to regulatory agencies or other stakeholders to demonstrate program implementation or effectiveness.

**Commented [BJ5]:** Note to EPA: Do you think we need to include Program Assessment – Using a combination of methods, an analysis of the overall effectiveness of the MS4 program? We have left it out for now as there are too many terms.

The definitions above use terms such as "activity" and "BMP," which also warrant being described to aid consistency in this report and ongoing dialogues.

- **Activity** – An action taken by a permittee or a regulated entity within the permittees jurisdiction that may provide a water quality benefit. Examples could be public education and outreach activities at a community event or outfall inspections.
- **Best management practice** (BMP) – A specific structural or non-structural management practice that is known to provide a water quality benefit. Examples could be a stormwater retention pond at existing development or erosion and sediment controls at a construction site.

## 2.4 Defining "Effectiveness"

Another purpose of the workshop was to explore the concept of effectiveness and how MS4 programs and permits can be improved to orient monitoring, tracking, evaluation, and reporting towards demonstrating effectiveness. This has proven particularly difficult and remains as one of the largest problems facing the national MS4 program. MS4 programs are often inherently complex due to multiple reasons—large geographic areas, numerous pollutant sources, transport of flows above and below ground in natural and manmade systems—and "despite these and other challenges, stormwater program managers find themselves facing increasing pressure to demonstrate the effectiveness of their programs" (California Stormwater Quality Association [CASQA], 2015). Further,

"**Effectiveness Assessment** consists of the methods and activities that managers use to evaluate how well their programs are working and to identify modifications necessary to improve results" (CASQA, 2015).

local programs' ability to carry out MS4 program requirements is often resource-constrained, thus making it increasingly vital to prioritize activities with outcomes that serve the community and environment. Many MS4 permits require local programs to evaluate the effectiveness of their efforts; however, there is tremendous ambiguity around what "effectiveness" means for MS4 programs—is it a measurable water quality outcome, completion of required activities, achievement of other co-benefits[2] through infrastructure improvements, or a combination of these?

Participants at the December 2017 and March 2018 workshops were asked to describe the key elements of MS4 program effectiveness, and though common themes emerged, there was significant variation in the responses. Based on responses from the pre-workshop questionnaires, key elements include:

- A clear definition of performance metrics (or measures) using common objectives and concise language.
- Impacts such as enhanced awareness and behavioral change.
- Reduction in urban stormwater pollution and mitigation of the impact on receiving waters.
- Tracking progress to ensure accountability of outcomes.
- An ability to measure and communicate quantifiable outcomes and benefits to communities.

The following are select responses from the pre-workshop questionnaire describing **key elements of MS4 program effectiveness**:

> *"Clear and measurable performance metrics and the ability to gauge activities and actions versus those metrics."*
>
> *"Ability to establish a relationship between the BMP/action/activity and a reduction in pollutant loads."*
>
> *"Ability to show water quality improvement, behavior change, and an overall understanding of the benefits and challenges associated with urban stormwater."*

Given the variation in responses and known difficulty in defining effectiveness, this paper does not attempt to create a single definition nor does it suggest that a single definition is feasible or needed. Rather, defining and determining effectiveness should occur at the permit, local, or regional scale and based on the unique conditions, objectives, and resources of the area. Throughout this report, the authors highlight various and situationally unique definitions of effectiveness. These are provided to demonstrate the various applications of effectiveness within the context of MS4 programs and how monitoring, tracking, evaluation, and reporting could be improved to facilitate a determination of effectiveness.

Workshop participants also wanted to address a common misconception that effectiveness is, in most cases, different than compliance. For example, a MS4 program could be compliant but may not be effective in addressing local water quality conditions, other co-benefits, or objectives.

[2] For example, reduction in flood risk, improvement in urban aesthetics and amenities through the use of green infrastructure, and water supply augmentation.

# 3 Overview of Monitoring Approaches and Efforts

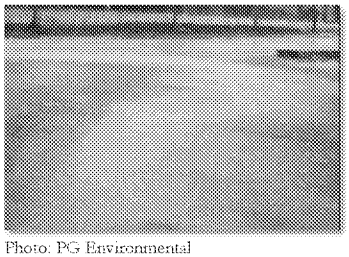


Photo: PG Environmental

Currently, permittees and regulatory agencies frequently evaluate program effectiveness through a combination of monitoring, tracking, evaluation, and reporting efforts, and other indicators). Though these requirements in MS4 permits are intended to enable iterative improvements, many programs do not systematically use their data in this way.

Stakeholders at the March 2018 workshop were asked to consider improvements to monitoring, tracking, evaluation, and reporting. For example, current approaches may not be optimized to detect change and correlate MS4 program actions with outcomes. There are also many examples where new permit requirements have been added to existing permits without careful consideration of how they enable performance evaluation and adjustment. This has increased the resources needed for monitoring, tracking, and/or reporting efforts and resulted in often lengthy and intense reporting efforts for permittees with little perceived benefit to the permittee, regulator, or water quality.

During the workshop, a Phase I permittee representative described how one of its MS4 annual reports filled 18 file boxes when printed. Permittees and regulators alike acknowledged the immense effort often expended by permittees on annual reporting and a common lack of resources at regulatory agencies to fully review and interpret submitted materials.

It is important to acknowledge, however, that there are diverse views on the need to improve stormwater monitoring. Some believe their MS4 programs are stable and reasonably effective; therefore, they don't require significant change, particularly as it relates to monitoring and evaluation.

## 3.1 Variation in Approaches

The national MS4 program was rolled out in two phases—Phase I targeted medium and large communities and industrial facilities, while Phase II addressed smaller communities and other non-municipal entities.[3] Both Phase I and Phase II regulations require permittees to assess their stormwater control measures (i.e., BMPs) and perform some level of reporting to regulatory authorities. However, variability within the regulations has led to significant variation in the way monitoring, tracking, evaluation, and reporting requirements are represented in permits and subsequently carried out by permittees. For example, Phase I regulations require permittees to develop a monitoring program, and larger MS4s may have requirements that necessitate sophisticated sampling programs with annual expenditures of over $1M. While the Phase II program allows for monitoring, it does not require it; as a result, some MS4 permits may not have any monitoring program at all (EPA, 2010a).

EPA's report from the 2017 workshop evaluating the MS4 program provides background on changes over time. For context, a workshop participant has characterized how one state's program

[3] For additional background on the MS4 program, see *Evolution of Stormwater Permitting and Program Implementation Approaches* (EPA Region 9, 2018).

has evolved. This is intended to provide a general point of orientation to provoke thought and further discussion. It does not attempt to capture the status or changes in all programs nationwide.

For some larger Phase I programs in California, there has been a general shift toward the inclusion of more water quality-based requirements with a focus on total maximum daily load (TMDL) implementation supported by modeling approaches. Permittees are capturing more program implementation data electronically through asset management software and other tools, and conducting water quality monitoring at various scales. However, there remains a need to improve monitoring, tracking, evaluation, and reporting approaches to better determine the effectiveness of program actions and allow for adaptive management over time. Observations on the overall conditions, approaches, and lessons learned in California's MS4 program are presented by era in Table 1 on the following page.

As noted above and in Table 1, all permittees must perform some level of reporting on MS4 program implementation activities for regulatory agencies. Therefore, tracking is necessary.

The following presents several examples of monitoring approaches used in different parts of the country.

- Through its principal permittee and a regional monitoring group, **Los Angeles County MS4 permittees** conduct monitoring in receiving waters and in-system locations for some design storms. Cause and effect connections are inferred to actions taken in the monitored watershed. Modeling using BMP effectiveness estimates for existing BMPs and accounting for anticipated load reductions for new BMPs is also used to estimate the likely overall effect of BMP implementation within watersheds.
- **Minnesota Phase II MS4 permittees** are encouraged to focus on implementation of minimum measures and not required to conduct monitoring. The State of Minnesota administers a statewide surface water monitoring program funded by a voter-supported measure, and there is an assumed correlation between MS4 program implementation actions and water quality effects.
- **Washington, D.C.**, has used geographically targeted monitoring designed to detect "signals" in water quality change based on intensive implementation of green infrastructure BMPs in the targeted area. Information gained at the smaller scale will then be extrapolated to evaluate larger scale implementation. The efforts include interim measurable milestones so the evaluation timeline is constrained.
- The **City of Salinas, CA**, in the Central Coast region has experienced an evolution of approaches since 2005 when there was a weak connection between water quality monitoring and program effectiveness. The program has moved from trying to assess the effectiveness of different program activities to focus more on structural BMP assessment and outfall load-based monitoring at several locations. This effort has been coupled with a web-based dashboard for tracking progress and offering information availability to regulators.

## Table 1 - The Evolution of California's Phase I Program

| Era | Early Generation Permits (1990s) | Middle Generation Permits (2000s) | Recent Generation Permits (2010s) |
|---|---|---|---|
| Overall Conditions | • Programs had limited knowledge of system assets and there were few known water quality drivers to direct program implementation. | • System assets were better known and there was increasing awareness of the need to address specific water quality issues (often through TMDLs) and to begin iterative program improvements.<br>• Newer data management tools were starting to be used and some information on BMP effectiveness was becoming available. | • There is a greater focus on specific POCs largely driven by TMDL provisions in MS4 permits).<br>• There are new automated and sensor-based monitoring methods that can enable different monitoring designs.<br>• There is an increased concern about asset management and long-term maintenance of system assets.<br>• There is a broader focus on stormwater impacts and value beyond water quality (e.g., water supply augmentation, flood risk, urban amenities/climate impacts). |
| Approaches | • Monitoring efforts were mainly focused on characterizing flows from the system and establishing baseline monitoring data for urban water quality conditions and trends.<br>• Sampling was required for a few storms per year, with little to no sampling during dry weather. Sampling was rarely conducted from MS4 outfalls; instead it was collected at convenient locations in the lower parts of watersheds to characterize "mass emissions" from all upstream MS4 discharges (often comingled with other sources and infiltration).<br>• Permittees typically designed their own monitoring programs.<br>• Paper reporting of water quality data and other program activity measures (e.g., inspections, street sweeping) through qualitative descriptions and/or semi-quantitative information | • More elaborative minimum control measure (MCM) requirements and narrative requirements to meet water quality standards (WQS) were included in permits. Some permits included numeric triggers or action levels for POCs and requirements for low impact development approaches for new/redevelopment.<br>• Some permits began to use surrogate indicators (e.g., flow retention, impervious cover) to reduce flows and pollutant loadings and protect receiving waters from geomorphic impacts.<br>• Monitoring efforts were mainly focused on receiving waters (rarely outfalls) to determine whether WQS were being met and whether MS4s were causing or contributing to exceedances.<br>• Permittees continued visual inspections of assets, BMPs, and dry weather flows and documented their occurrence in annual reports. There was rarely an analysis of their effectiveness. | • Permits include more specific water quality-based requirements, often connected to TMDLs.<br>• Permit structure varies depending upon whether goals are expressed in terms of outcomes (numeric limits or triggers) or activities (BMP systems based on analysis of needs).<br>• Models are increasingly used to inform long-term program design and predict necessary control levels.<br>• Minimum control measures remain but, in some instances, focus on a subset that are viewed as more effective.<br>• Adoption of asset management allows for operations and maintenance activity reporting and a determination of optimal asset inspection and maintenance schedules.<br>• There is less of a focus on basic water quality trend monitoring in receiving waters and more of a focus on representative outfall monitoring to help evaluate causation.<br>• There is an increasing use of automated samplers, but in limited locations. |

| Era | Early Generation Permits (1990s) | Middle Generation Permits (2000s) | Recent Generation Permits (2010s) |
|---|---|---|---|
| Lessons Learned | • Monitoring program designs rarely enabled key management questions (including compliance questions) to be answered based on the collected data.<br>• Insufficient data was collected to detect pollutant trends in receiving water or distinguish among contributing land uses or geographical areas.<br>• There was insufficient evaluation and reporting to ensure that stormwater controls (e.g., post-construction BMPs) were installed and properly maintained.<br>• There was insufficient data or analysis to evaluate effectiveness of MCMs or other activities/BMPs in addressing specific water quality concerns. | • There was still insufficient data collected to detect pollutant trends in receiving water or distinguish relative contributions from different land uses, geographical areas, or individual permittees.<br>• Monitoring designs did not support robust statistical analysis or provide a linkage between receiving water impacts and specific MS4 discharges (i.e., unable to answer the key question of whether the MS4 was causing or contributing to a WQS exceedance).<br>• There was still insufficient data and analysis regarding BMP effectiveness to determine whether installed BMPs were resulting in the intended benefits.<br>• Reporting and program evaluation still did not thoroughly address the effectiveness of MS4 programs in creating the desired water quality outcomes. | • Improvement is still needed to evaluate the effectiveness of activities performed under the MCMs.<br>• In many cases, program implementation and monitoring requirements continue to mount while few are removed from permits.<br>• New sensor technologies are not widely being used in monitoring program design.<br>• Much receiving water and outfall monitoring still does not facilitate source analysis, compliance evaluation, or effectiveness evaluation.<br>• There is a need to better understand how increasing reliance on modeling affects monitoring and reporting needs.<br>• Modeling capacity and monitoring design will need to evolve to better account for non-water quality intended benefits (e.g., water supply augmentation through infiltration, reduced flood potential, heat island impact reductions). |

## 3.2 Existing Efforts and Resources

Since the MS4 program began, several entities have articulated potential improvements for program monitoring, tracking, evaluation, and reporting approaches. Despite these efforts, there is still a clear need for a concise and reproducible approach to monitoring and evaluation which yields actionable information with linkages to water quality outcomes. Further, additional training for permit writers and permittees is needed to build overall capacity relating to monitoring and evaluation strategies. The following identifies select examples of existing resources; it is not intended to be fully comprehensive.

**EPA Region 3**'s *Evaluating the Effectiveness of Municipal Stormwater Programs* describes a process of goal setting in stormwater management programs (SWMPs), matching evaluation to management goals, evaluating SWMP effectiveness through a combination of program operations (e.g., activities), social indicators, and water quality monitoring. The document excerpt below displays an example thought process of matching evaluation to management goals and the corresponding actions needed to measure and assess.

> "Evaluation of the effectiveness of a SWMP must relate directly to its goals. Two central questions are: *Are we meeting the municipal SWMP goals?* and *Are we meeting NPDES stormwater regulatory requirements?* If a goal is to keep a swimming beach open, it is often necessary to determine the extent to which water quality criteria for bacteria are being met. If a goal is to reduce nutrient loads by 40% from a watershed, it is then necessary to measure nutrient loads and compare measured loads against the goal. Meeting your water quality goals is the ultimate sign of program success, however, meeting programmatic or social goals can also be indicators of a successful program. Information on how these goals are met will serve as critical feedback in the iterative process of stormwater management." (EPA, 2008)

The **California Stormwater Quality Association** (CASQA) has also done significant work related to MS4 program effectiveness assessment and monitoring since the early 2000s. CASQA's more recent guide titled *A Strategic Approach to Planning for and Assessing the Effectiveness of Stormwater Programs* is a comprehensive 500-page reference intended to "establish specific 'how to' guidance with examples for managers in planning and assessing their MS4 programs" (CASQA, 2015). The document introduces the concepts of six key outcome levels that provide "structure and measurability to evaluate and improve Stormwater Management Programs over time." The outcome levels (depicted in Figure 2 below) provide a basis for discussion of how progress can be measured for MS4 program elements through monitoring or other means. This is an important resource to consider while developing a vision for the future of stormwater monitoring to improve program efficiency and effectiveness. CASQA also developed a *Program Effectiveness Assessment and Improvement Plan Framework*, an approach and format for permittees to assess and document MS4 program effectiveness that based on their guidance document. Many MS4 permittees in California are required to use this, or a modified process, to perform effectiveness assessments.

**Commented [AK6]:** Note to EPA: Need to seek permission from CASQA.

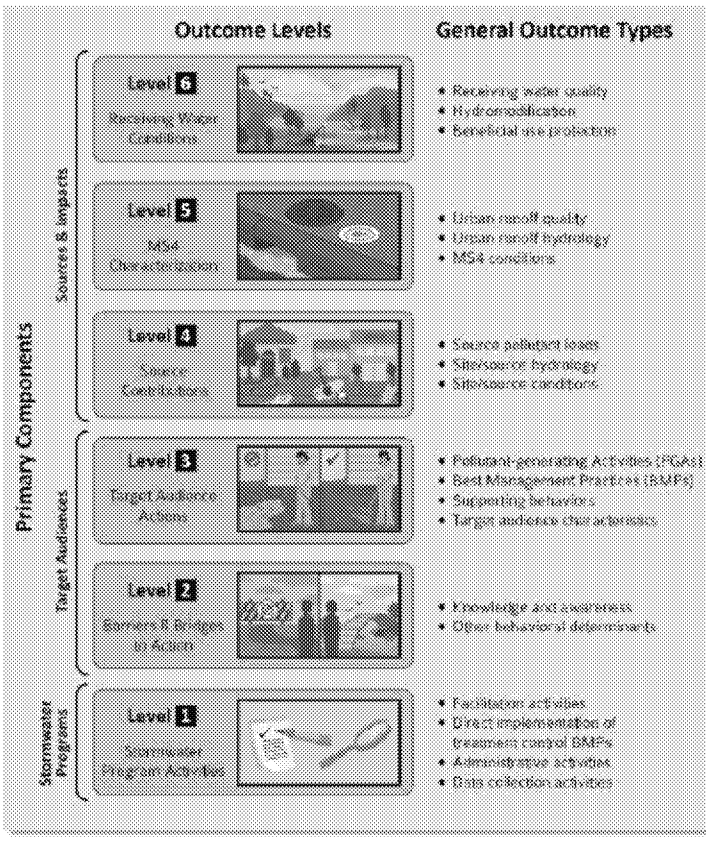



**Figure 2.** CASQA stormwater quality outcome levels. Graphic included in this report with permission from CASQA (CASQA, 2015).

The **Southern California Stormwater Monitoring Coalition** has developed several reports and resources for its members to "better understand stormwater mechanisms and impacts" and "develop the tools that will effectively and efficiently improve stormwater decision-making" (SMC, n.d).

**EPA's** *MS4 Program Evaluation Guidance* is a guidance document developed for state and NPDES permitting authority staff to assess compliance and effectiveness of MS4 programs. This document has served as the basis for compliance audits since its publication. The document notes that "the findings of the MS4 evaluation should not be based solely on the level of achievement of measurable goals. It is important, however, that the permittee's SWMP includes the use of measures to assess progress towards meeting goals that benefit water quality and not rely on 'bean-counting'" (EPA, 2007).

**The California State Water Resource Control Board**'s *Guidance for Assessing the Effectiveness of Municipal Storm Water Programs and Permits* was developed to assist State Regional Water Board staff in assessing the effectiveness of the storm water programs implemented by local agencies. The document incorporates CASQA's outcome levels in its process and "lays out a framework for assessing the effectiveness of MS4 program implementation as a whole, rather than looking at the individual programmatic elements" (CASWRCB, 2010).

**The Center for Watershed Protection**'s document titled *Monitoring to Demonstrate Environmental Results: Guidance to Develop Local Stormwater Monitoring Studies Using Six Example Study Designs* presents monitoring study designs to help communities develop monitoring studies that will improve local stormwater programs (CWP, 2008).

# 4 Recommended Program and Permit Improvements

During the workshop, facilitators encouraged participants to identify tangible ways to enhance program implementation and permit efficiency and effectiveness to protect water quality. These conversations generated a wide range of recommendations related to monitoring, tracking, evaluation, and reporting under the following broad headings:

Photo: EPA

Commented [WG7]: Note to EPA: This appears to be a central purpose of the workshop. Do you agree? Are there others? We should try to be specific about the purpose(s).

- **Recommendations for Capacity Building and Program Support** (Section 4.1)
- **Permitting Recommendations** (Section 4.2)
- **Making Outfall and Receiving Water Monitoring More Discriminating to Inform Program Management** (Section 4.3)
- **Improving Our Ability to Detect Effectiveness—Approaches to Link Water Quality Outcomes to Actions** (Section 4.4)
- **Improving Program Tracking and Reporting** (Section 4.5)

The set of recommendations presented in this report is not definitive nor is it exhaustive; rather, this report is intended to serve as an inspiration for further discussions and follow-on actions. References to select projects or organizations are incorporated throughout to serve as case studies and examples of related efforts.

## *4.1 Recommendations for Capacity Building and Program Support*

While approaches to monitoring, tracking, evaluation, and reporting can be viewed in their own lanes, they are intrinsically linked together and, to some degree, need to be considered collectively to identify meaningful improvements. As such, workshop discussions often focused on this holistic view and resulted in several overall recommendations to build capacity related to monitoring, tracking, evaluation, and reporting. Collectively, the following strategies could improve overall MS4 program effectiveness and water quality outcomes.

### 4.1.1 Develop a Vision for the Future of Stormwater Monitoring to Improve Program Efficiency and Effectiveness

During the workshop, participants identified a lack of a central vision for why local programs perform monitoring—what questions do we need to answer now and into the future—and how these efforts relate to program evaluation, tracking, and reporting. Participants highlighted significant inefficiencies in how these activities are typically carried out and noted potential for improvement with benefits for local programs, regulators, and water quality. Notably, many participants expressed concern that many municipal stormwater monitoring efforts are resource-intensive and yield little actionable information for management decisions. Some participants also emphasized an acute need for models to enhance program capabilities for planning and program assessment; otherwise water quality monitoring across large geographic areas and time scales can be

resource-prohibitive. However, along with increased model usage comes a need for increased water quality monitoring data to inform and validate models.

During the workshop participants discussed what they envisioned to be key attributes of a more effective approach to monitoring and how it may intersect with other evaluation, tracking, and reporting efforts.

- Clear management questions related to water quality outcomes and activity implementation.
- A process for conducting effectiveness assessment that is tailored to the program element and the management questions being asked.
- Use of improved monitoring designs (location, scale, frequency, methods) to detect a "signal" or change in pollutant loading in stormwater or receiving waters for POCs.
- Monitoring efforts that are complementary to and aligned with activity tracking and assessment to better evaluate effectiveness of structural or non-structural controls (e.g., are they implemented correctly, receiving proper maintenance, and operating as expected?) and improve the basis for assessing cause and effect.
- Documented monitoring and evaluation designs coupled with identification of program modifications envisioned to improve effectiveness and achieve intended outcomes.

As noted above, workshop participants identified pollution reduction, water quality protection/improvement, enhanced awareness, and behavior changes as some key elements of program effectiveness. To achieve these outcomes and guide program implementation, workshop participants also noted a need for clear program performance metrics (further described below in Section 4.1.3).

### 4.1.2 Develop Guide to Improving Monitoring and Evaluation to Better Serve MS4 Programs

Commented [AK8]: Note to EPA: Should this be a guide on Effectiveness Evaluation? Or just Monitoring? Or the whole enchilada? How monitoring can be used as part of effectiveness evaluation?

Just like we had terminology issues in the first report, I think we have them here a bit with the terms evaluation, effectiveness, assessment, and how sometimes those terms get put together – program effectiveness assessment, effectiveness evaluation, evaluating effectiveness, assessing effectiveness, etc. – let's discuss how to address in the next version.

**100 percent of** pre-workshop questionnaire respondents agreed that *"Monitoring designs must go beyond just data collection methods to include data management, data analysis, and reporting formats that clearly link data collected with performance metrics."*

Currently there are various approaches to monitoring and evaluation used across the country. As described above, some involve a state-run surface water monitoring program with a certain level of association with local MS4 programs, others involve a mix of receiving water and outfall monitoring and activity tracking and evaluation, and yet others are implemented in smaller watersheds to evaluate the effectiveness of specific types of stormwater control and treatment practices.

Workshop participants discussed a need to identify the range of monitoring approaches used and how they associate cause and effect (i.e., are MS4 program actions impacting water quality conditions). Further, some participants suggested using this effort to identify successful designs to inform a national level guide on monitoring and assessing program effectiveness. This could promote consistency across the national MS4 program and enhance efficiency in local program implementation and efforts by regulators during permit development and compliance review. For example, one participant indicated that 34 stormwater monitoring groups in southern California were unable to develop common monitoring questions due to differences in study designs, methods, or data management systems.

The proposed guide could be informed by existing resources on monitoring and effectiveness and the entities involved in their development and ongoing monitoring design efforts. Workshop participants suggested this guide should include the following elements:

- Framing monitoring/evaluation questions and designing approaches to fit the questions. Specifically, this could include alternative program designs with advice on assembling the components (e.g., receiving water, outfall, and in-system water monitoring; BMP effectiveness monitoring; activity tracking of structural and non-structural controls; modeling) to demonstrate effectiveness. This should show how to build a sound analytical framework up front to demonstrate why a set of approaches will likely be successful in assisting program management and defining or tracking compliance and effectiveness.
- Considerations for adapting monitoring/evaluation questions over time with a reasonable limit to the creation of new questions.
- Examples of successful local approaches that better associate monitoring/evaluation design with program effectiveness, compliance assessment, and the ability for program managers to make management decisions.
- Available monitoring technologies and best practices that clearly link the monitoring objectives with the experimental design, including all aspects of data collection, data management, data analysis, and reporting formats.
- Compiling monitoring program costs to help show the wide range of program expenditures, how monitoring data is used to inform program decisions, and how to better articulate the value of the data.
- Explanations of modeling approaches and how they can relate to monitoring and adaptive management.

Beyond the monitoring design elements, select workshop participants suggested that clearer direction is needed for the technical aspects of monitoring as well. Specifically, standard protocols and references are needed appropriate equipment, protocols, site selection, sampling frequency, data management/analysis, and quality assurance. Program evaluators (e.g., regulators) also need guidance in assessing the technical "quality" of discrete monitoring program elements.

**2NFORM** is an ESRI mapping program specific to stormwater. It can be used to design a monitoring approach and help with tracking/analyzing/reporting. It is also capable of linking to sensors/monitors from the field.

For their guidance on mapping capabilities, see: http://www.2ndnaturellc.com/documents/MS4_Mapping_Guidance.pdf

### 4.1.3 Establish Key Performance Metrics (Activity- and Outcome-Based) for Municipal Stormwater Programs

There was agreement among workshop participants that clear performance metrics need to be established to enable meaningful MS4 program evaluation and monitoring efforts. Participants discussed ideas for developing metrics that are valuable and can help define measurable outcomes. Multiple people suggested that efforts are needed to compile possible metrics (from prior efforts such as rulemakings or new metrics) and synthesize the information to help progress in this area.

**Ninety-six percent** of pre-workshop questionnaire respondents agreed that *"Performance metrics need to be established in concert with improved monitoring designs and methods"* and that *"Metrics should enable evaluation not just of what was done, but also of whether those actions were effective."*

It was noted, however, that it may not be possible to identify meaningful performance metrics with measurable outcomes for some MCM activities. Further participants indicated that there should be specific considerations for the differences between structural BMPs (e.g., permanent stormwater controls) and non-structural BMPs (e.g., facility inspections).

During a facilitated exercise, workshop participants brainstormed possible overall metrics as indicators of program performance that go beyond tallying activities or "bean counting." Below is a list of ideas put forth by participants.

- Percent of impervious areas addressed for stormwater management.
- Condition or "cleanliness" of streets as an indicator of potential pollution from runoff.
- Percent of parcels connected to the storm drain system.
- Modeled volume of flow to the storm drain system used as a surrogate for pollutant contributions.
- Percent of waterbodies in a community that are fishable and swimmable.
- Loss of beneficial use of a waterbody (e.g., beach closure downtimes).
- Measured level of awareness of citizens regarding stormwater pollution and the community's program.
- Increasing number of illicit discharges reported annually; indicating heightened awareness.
- Budget for stormwater infrastructure improvements.

The **American Water Works Association** has a benchmarking program for drinking water programs; no analogous program exists currently in the stormwater sector.

The **National Municipal Stormwater Alliance** (NMSA) is currently working with the **American Society of Civil Engineers** to develop a national stormwater "report card" since data on program performance is lacking.

Participants also discussed several MCMs and whether clear links could be drawn between program activities and measurable water quality outcomes. It was easier to envision linkages for water-quality based effort such as requirements for stormwater management in new development and redevelopment, while activities like public education and outreach, construction site inspections, outfall screening activities proved more challenging. One workshop participant characterized it this way: "There is an obvious desire to seek and set *outcome* rather than *output* performance metrics. However, MCMs are primarily or essentially low-cost prevention actions, which don't lend themselves to measurable water quality outcomes."

Questions remain as to what are meaningful performance metrics for MS4 programs overall as well as the individual program elements and MCMs. Further, what is an appropriate mix of output and outcome metrics that can guide programs in developing monitoring programs, assessing effectiveness, and performing tracking and reporting functions?

### 4.1.4 Leverage Existing Data Sets to Enhance Approaches for Informing Program Management Decisions and Gauging Program Effectiveness

MS4 programs have collected, documented, and reported a significant volume of data on implementation and monitoring over the years. While some permittee representatives at the workshop lamented the amount of resources typically involved in tracking and reporting, they also acknowledged that the vast amount of data collected has the potential to inform program management decisions.

Workshop participants suggested that better data analytics tools, processes, and guidance need to be developed for program managers to (1) turn existing data into information, (2) use the information to more confidently make program management changes, and (3) collect better data to continue to feed the process. For example, existing data sets regarding illicit discharges could be analyzed in concert with outreach information and awareness levels to identify trends and better direct program resources. One inherent issue is that local programs use various mechanisms for tracking data and not all programs track the same types of data. This issue will need to be considered and addressed, and the development of new tools with tangible uses could encourage more consistency in data collection techniques.

There was also discussion at the workshop regarding the possible use of data in annual reports submitted by local programs in a state or region. Workshop participants noted that many NPDES permitting authorities do not have resources to fully review the significant quantity and volume of annual reports submitted; however, within those reports there may be some intermediate indicators of program performance that could readily be identified to provide feedback to permittees. Further, trends observed in a group of annual reports in a state or region could be used to inform permittees of common issues and areas requiring more clarification or support to yield better program implementation. Many states or regions have municipal stormwater management groups that meet periodically and could serve as a forum for sharing this type of information.

California's **Storm Water Multiple Applications and Report Tracking System** is a web-based platform for stormwater program (construction, industrial, municipal) permit applications and reporting. Workshop participants suggested that data in this system could be used to help inform some municipal stormwater program functions and priorities, especially as it relates to oversight of construction sites.

## 4.2 Permitting Recommendations

**Eighty-eight percent** of pre-workshop questionnaire respondents agreed that *"Stormwater quality monitoring has been largely ineffective in assisting compliance evaluation, problem targeting, and program improvement."*

As noted above, many permitting authorities and permit holders believe there are significant opportunities to improve approaches to municipal stormwater program monitoring, tracking, evaluation, and reporting, and these improvements may be directed or even incentivized through permitting strategies. Workshop participants generally recognized that permit writers may not have available tools or guidance to craft better monitoring, tracking, evaluation, and reporting requirements and may not be willing to stray from the status quo.

As was also noted in the first workshop, MS4 permitting programs are often understaffed and have devoted insufficient resources to provide technical and policy guidance, assist permittees in program improvement, and issue timely permitting decisions and compliance actions. Provision of adequate resources for EPA and state permitting offices will be critical to facilitating improvements in permitting and program development.

### 4.2.1 Improve Clarity of Monitoring and Effectiveness Permit Requirements and Objectives and Methods/Designs

Workshop participants suggested that current permit designs for monitoring requirements often lead to long-term monitoring at geographic scales which do not enable detection of signals for program performance or establish cause and effect between program actions and water quality conditions. Further, there is often an aversion to modifying monitoring locations for fear of losing continuity in long-standing datasets. Some of this may be due to ambiguity in permit requirements or reluctance by permit writers to change requirements; it may also be an attempt to put the onus on permittees to develop monitoring programs without guidance to steer them toward more efficient and effective designs. Many permits also give equal weight to tracking and reporting for all aspects of program implementation, which can lead to incredibly resource-intensive efforts by local programs to record, compile, and summarize information for annual reporting efforts.

**Ninety-two percent** of pre-workshop questionnaire respondents agreed that *"Training and outreach for permit writers, program staff, and elected officials on new methods and designs are needed to familiarize these groups with their benefits and limitations."*

Workshop participants identified an overall need for permitting authorities to improve the clarity of monitoring and effectiveness permit requirements and to use thoughtful methods/designs that will yield actionable data. Further, some participants noted that permits may be able to provide choices or flexibility for monitoring approaches and help incentivize better designs.

To help permitting authorities to understand various approaches being used across the country, EPA developed a compendium series of MS4 permitting approaches. Part 3 focused on water quality-based requirements and included a section describing monitoring and modeling approaches related to TMDLs and water body impairments. While this compiled information is helpful for understanding some relevant monitoring case studies, it does not evaluate what aspects of the efforts were successful or unsuccessful, identify benefits and limitations, or provide advice on what approaches are appropriate for certain scenarios. There is a continuing need to identify different approaches and extract the lessons learned to inform efforts by others in the sector.

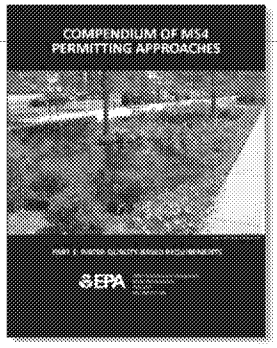



**Figure 3.** Water quality-based requirements compendium that includes several monitoring program examples.

**Commented [BJ9]:** Note to EPA: Karen suggested that CASQA's PEAIP framework presents a method for building programs around pollutants of concern. Do you think that would be an example worth exploring or mentioning here or elsewhere in the report?

California's Municipal Regional Stormwater NPDES Permit for the San Francisco Bay area (adopted 2015) features a monitoring program that is driven by management questions, allows for scaling up to larger areas (county-wide or region-wide), accounts for different types of monitoring (e.g., receiving water status monitoring, POC monitoring), and includes stressor/source identification projects in response to monitoring findings. The permit provides directions on various methods to obtain relevant information to drive management actions. The monitoring requirements have attempted to provide a balance between directives and flexibility to allow permittees to seek optimum benefit from monitoring with available monitoring resources.

It should be noted that increased clarity and better designs may not be possible to achieve without first accomplishing some of the recommendations for capacity building and program support described above. Training and other support tools will be needed to help boost permit writers' understanding and ability to improve approaches to monitoring, tracking, evaluation, and reporting.

### 4.2.2 Create a Pathway in Permits to Make Special Studies More Impactful

Special studies or additional monitoring requirements are often included in NPDES permits to help gather data needed to explore identified issues and support future permit development. The NDPES Permit Writers' Manual notes that permit writers should establish reasonable schedules for completion and include in the permit any requirements (e.g., special sampling, analytical procedures) related to the study (EPA, 2010b).

Workshop participants indicated that, especially in California, there are many long-term or special studies completed, but there often isn't the opportunity to apply the lessons learned from the efforts. Participants urged that *if* special studies are required, there should be a clear pathway in the permit to apply the lessons learned. Further, some participants noted that special studies should be designed to address a specific topic and result in a short-term study with a discernible beginning, middle, and end—a process to obtain the answer to the question, apply the knowledge, and then move on.

Some participants described special studies as an opportunity to be more targeted in scope. In such cases it would not necessarily have to relate to overall program effectiveness, rather it could be used to improve program operations. For example, special studies could be a testing ground for exploring the use of innovative technologies, sensors and screening devices, or remote sensing on smaller scales before a program makes a significant investment and a permit writer moves any associated requirements into the core permit. There could be a tiered approach that links the research field to the regulatory community to help field test new technologies.

Commented [BJ10]: Note to EPA: In the workshop notes we have "SMC is a framework for stakeholder involvement for developing special studies – all of S Cal putting money into effort and engaged with various regional boards." Do you think this would be appropriate to reference here under special studies or elsewhere? Do you already know additional information to include?

One workshop participant put forth the following straw proposal of how special studies could more effectively be viewed within the construct of an overall monitoring approach.

- Special studies should explore very specific, complex questions. If the questions are answered, then the benefits could extend far into the overall MS4 program.
- Sophisticated equipment and protocols may be needed for special studies, though the outputs should be simple and applicable to help a program adapt.
- Not all permittees should be asked to perform special studies—there should be fewer, more specific special studies to answer questions facing the program.
- Some questions (e.g., BMP effectiveness) may not be appropriate to address through permits; outside parties should be engaged to help.

In summary, participants saw an opportunity to improve how the results of special studies are applied to the not only the programs that conducted the studies but, in some cases, the larger community of MS4 programs. However, at present, there is a gap in bringing this knowledge to the broader program.

### 4.2.3 Identify Instances When Lack of Approved Monitoring Methods Created a Barrier to Technology Implementation

There is a proliferation of new technologies for measuring water quality, with an increasing trend toward continuous, real-time sensors. Other "bio" technologies are being developed to detect the presence of certain parameters. Approved sampling and analysis methods at 40 Code of Federal Regulations (CFR) Part 136 do not necessarily include these new methods, which workshop participants identified as a potential barrier to the use of new technologies. Validation of new technologies was identified as a hindrance to both technology developers for commercialization and for program managers to confidently move forward with using a new technology.

As an action item, workshop participants suggested inventorying known instances of where programs have elected not to use a new monitoring technology because it is not an approved method. Where possible, it would be helpful to identify avenues to address identified issues, whether through rule changes or other creative uses of the technologies, to improve program operation.

Representatives from environmental organizations at the workshop explained that they often employ new technologies that are not approved by 40 CFR Part 136 as they are not beholden to permit-approved methods for their research activities. This group possibly represents a part of the sector that may be more willing to test new approaches and then share with the broader program.

Additional discussion about envisioning uses for sensors and other new technologies is included in Section 4.3.2 below.

## *4.3 Making Outfall and Receiving Water Monitoring More Discriminating to Inform Program Management*

Water quality monitoring is necessary in most cases to demonstrate NPDES permit compliance. In addition, results from outfall and receiving water monitoring can provide useful datasets and key indictors to inform stormwater program management and decision-making. Expanding the use of monitoring within a program and making it even more discriminating can lead to more effective implementation of the stormwater program as a whole.

Increasing data collection alone, of course, will not automatically lead to improved program effectiveness. One hundred percent of survey respondents agreed that monitoring must go beyond data collection to include data management and analysis that links the acquired information to specific performance metrics. Workshop participants indicated a need for guidance in designing monitoring programs to yield actionable results and for assistance in linking monitoring data to programmatic decision-making. They also expressed a desire to expand the use of real-time monitoring for stormwater operations and supported deploying pilot programs and special projects for innovative monitoring technologies coming to market.

### 4.3.1 Evaluate Options and Effectiveness of Conducting Monitoring Efforts at Scales Designed to Yield Actionable Results

Workshop participants described the importance of scale when conducting monitoring efforts. By first identifying problems or questions, programs can ensure appropriate geographical range and time scales in monitoring approaches to address them (e.g., large scale for broad regional/watershed perspective, small for studying specific areas for pollutant contributions/mitigation).

**Commented [S11]:** At some point we should have a separate text box describing how DC water is doing intensive implementation paired with intensive monitoring in small geo area of DC

**Commented [WG12R11]:** Note to EPA: We will work with Bethany to collect a more detailed description of her program for inclusion. It very well may end up being more than a text box. Possibly we could extract the critical elements for success in addition to a basic overview.

In pursuing monitoring efforts that incorporate multiple jurisdictions (e.g., regional, watershed, and statewide levels), each program will have to consider the value of sharing data with their surrounding communities. When program functions are shared through partial consolidation at watershed or regional scales, there may be opportunities for aligning monitoring, tracking, evaluating, and reporting activities.

However, there can also be hurdles in extracting and comparing data for large-scale monitoring efforts that comprise several jurisdictions. Therefore, MS4 programs should ask themselves several questions before embarking: Are there opportunities for resource savings over the long-term? Are there incentives that can be offered for integrating new jurisdictions into existing monitoring programs? Can sampling, analysis, and data management and interpretation be standardized to allow for comparability? Does collected data help to answer established management questions for each participating jurisdiction? These are just some of the considerations that need to be accounted for when weighing the pros and cons of increasing the scale of a monitoring effort. Even in cases where no formal partnering is established, workshop participants suggested that monitoring and annual reporting requirements can be structured to provide an opportunity for comparability, information sharing, and technology transfer within a state, region, or nationally.

Monitoring on small scales is equally important and workshop participants identify several small scale and targeted monitoring efforts that have produced tangible results linking program efforts to water quality improvement. One such example is provided in the text box regarding DC Water's

efforts. Participants thought successful small-scale efforts should be identified and included in guidance, case studies, or other means to inform future efforts and provide lessons learned.

### 4.3.2 Convene Visioning Session for Deploying Sensors in Municipal Stormwater Programs

Workshop participants suggested convening a visioning session focused on the identification acceptance, and deployment of sensors in municipal stormwater programs. Visioning topics should include the use of sensors for improving system operations (e.g., illicit discharge detection, pipe clogging, flooding) as well as for designing and implementing real-time control programs to better manage water resources. These topics align well with existing and ongoing work being done through EPA's Office of Water's water technology and innovation (e.g., "Intelligent Water"). The visioning sessions should acknowledge the barriers discussed previously in this report and present a range of remedies.

WRF's **Leaders in Innovation Forum for Technology** (LIFT) is a multi-pronged initiative to help bring new water technology to the field quickly and efficiently. Intelligent Water Systems has been selected as one of their key focus areas. Subscribers can participate in regular discussion forums and presentations on the topic, access technology evaluations, and review the latest research.

Workshop participants acknowledged a need for more impactful studies surrounding innovative technology, particularly for sensors and real-time controls. Further, there is a need for broader dissemination of information related to current technologies and best practices available for water quality monitoring. The visioning session could be used as a platform to identify additional opportunities for special projects for permit inclusion to pilot innovative technologies to improve water management and enhance decision making.

## *4.4 Improving Our Ability to Quantify Effectiveness—Approaches to Link Water Quality Outcomes to Actions*

Since MS4 program inception, many regulators have largely employed a "faith-based" approach for assessing program effectiveness related to water quality improvements. That is, if the components of a permit are implemented adequately, it is assumed that will lead to improved water quality. There was little or no data provided to support such conclusions. To date, very few programs have gone so far as to analyze and document the actual effectiveness of their programmatic measures and physical BMPs at removing pollution from stormwater runoff.

Photo: PG Environmental

Some permittees have established sophisticated monitoring and modeling to better quantify the effectiveness of their stormwater programs; however, for the majority, a realistic and effective approach for demonstrating the specific link between actions and water quality improvements has been elusive. This can be due to many factors. For example, watersheds

and drainage areas may be quite large with many small sub-drainage areas where stormwater is managed, either through targeted programmatic practices or physical BMP treatment. This can create the need for numerous upstream and downstream monitoring locations to accurately determine the effectiveness of implemented actions. Beyond the logistical hurdles a permittee may face, a widespread monitoring effort would likely be cost-prohibitive for the average permittee. Further, stormwater pollution sources are often dynamic (constantly changing) and vary widely.

**DC Water** described pre- and post-monitoring activities for two green infrastructure installations. The $1M cost, approximately 2 percent of the overall project budget, was funded through their impervious surface charge. A dedicated team oversaw sensor installations and ensured that equipment stayed in the system over a multi-year period.

Workshop participants agreed that the programs should move away from the "faith-based" approach and focus on improving capabilities for determining and quantifying the actual effectiveness of specific actions on improving water quality. There was an acknowledgement that useful data may exist that has not yet been tapped for this purpose (e.g., turbidity and sediment loss data for construction sites, data collected for rulemaking purposes). Likewise, transferable approaches have been deployed in other programs such as for combined sewer overflows. Workshop participants communicated the need for better tools, guidance, and methods for accurately quantifying BMP performance.

### 4.4.1 Document Current State of Knowledge of BMP Performance and Effectiveness

Workshop participants were divided in their assessment of the current state of knowledge on BMP performance and effectiveness. Some thought there was a robust cache of data available, while others saw a clear need for more and better information. In either case, there was an acknowledged need for improvement in publicizing the results of unique and beneficial datasets regarding BMP performance and effectiveness to promote better knowledge transfer.

During the previous workshop assessing the overall MS4 program, participants identified that performance of structural and non-structural BMPs[4] needs to be better measured and reported for existing approaches as well as new technologies as they come to market. The resultant report acknowledged available data and information are particularly limited concerning effectiveness of non-structural BMPs such as public education, illicit discharge controls, and facility inspections. These non-structural elements are the main building blocks of the traditional MS4 programs.

Some publicly-accessible resources do exist with documented examples of BMP performance data. For example, the International Stormwater BMP Database includes over 600 datasets, publications, and tools related to stormwater BMP effectiveness. The Database is well positioned to host and disseminate documented test results and studies from many of the leading organizations addressing the topic of BMP effectiveness, such as WEF and its National Stormwater Testing and Evaluation for Products and Practices (STEPP) Initiative, which is aimed at validating the performance of innovative stormwater management technologies. Other organizations, like CASQA are working at the state or regional level to develop more locally-focused tools to help quantify the water quality impact of stormwater program actions (e.g., calculating source-load reduction).

[4] In the stormwater program, there is often overlap and ambiguity in the terms used to describe practices to control the volume and/or quality of stormwater runoff (e.g., post-construction BMPs, permanent stormwater controls, structural BMPs, non-structural BMPs). For simplicity and consistency, this report uses "BMPs" to include these types of control measures in both gray and green infrastructure applications.

Workshop participants acknowledged that despite the currently-available resources, there is still a need for more research and information sharing to improve our ability to quantify the effectiveness of stormwater program actions. Broadly-inclusive databases can be a good starting point, but additional data reflecting location-specific information such as geomorphology, hydrology, climate, O&M strategy, and the presence of unique or emerging pollutants is needed. Ultimately, increasing the variety and robustness of data and information about different BMPs' performance and effectiveness is needed to build the capacity of local programs, public agencies, and private parties to implement the most-appropriate methods for specific pollutants under local conditions.

### 4.4.2 Improve the Applicability and Usefulness of Modeling through Collecting and Incorporating Better BMP Performance Data

Modeling is and will likely remain a primary method for estimating stormwater program effectiveness and BMP performance. While no model will ever be 100 percent accurate, they can become more useful through high-quality data that is representative of the real-life conditions. Workshop participants expressed concern that the current limitations in effectiveness and performance data have resulted in lower confidence in the ability of models to be useful across a wide variety of stormwater management settings (e.g., different regions, climates, hydrology, geomorphology).

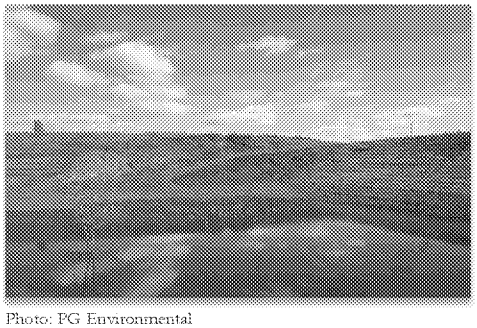


Photo: PG Environmental

Workshop participants acknowledged the need to collect more and better effectiveness data for all BMPs to improve the usefulness of modeling, especially for non-structural BMPs (e.g., public education and outreach, illicit discharge detection and elimination, facility inspections). Non-structural BMPs can be a critical for reducing runoff pollution, but they are often left out of stormwater models because their effectiveness is difficult to quantify and there is limited data available on these practices. Proactive and preventative pollutant source control methods such as street sweeping and other good housekeeping measures are also underrepresented. In stormwater modeling, the effectiveness of BMPs has traditionally been calculated based on runoff volume reduction (i.e., pollutants are reduced through decreasing the volume of runoff carrying those pollutants).

EPA's Storm Water Management Model (SWMM) is a robust tool used worldwide to estimate the effects of stormwater runoff on collection systems and the environment. SWMM conducts hydraulic and hydrologic simulations and has the capacity to estimate pollution reductions related to BMP implementation (EPA, 2016). EPA's National Stormwater Calculator helps developers assess the impacts of runoff from the impervious surfaces on their projects. It also provides guidance and runoff reduction estimates that can be used to help select effective low impact development controls (EPA, 2017). Though powerful, functional limitations include the need to be broadly-applicable as well as the limited set of basic structural practices. In both cases BMP effectiveness estimates are driven by runoff volume reduction achieved through structural BMP implementation.

Ultimately, stormwater managers need useful models that inform decisions and quantify progress simultaneously. This requires a mechanism with the ability to utilize all factors contributing to pollutant reduction and incorporate new information and adapt model outputs over time.

### 4.4.3 Establish Model Calibration Guidance to Improve the Usefulness of Models

Commented [WG13]: Note to EPA: 4.4.2 and 4.4.3 can likely be combined. Additionally, this appears to be a good location to introduce and describe the RAA.

Models are predictive tools that become increasingly useful as better and more applicable data is input for calibration. Current stormwater management models, like SWMM, often rely on parameters focused on reducing runoff volume and rate (EPA, 2016). In SWMM's case, calibration considers parameters related to imperviousness, storage capacity, and the physical characteristics of conveyances (Barco, Wong, & Stenstrom, 2008). Models have long been used to aid in planning for development, but workshop participants acknowledged models are increasingly being used to supplement water quality monitoring and provide flexibility to permittees when a widespread comprehensive monitoring program is infeasible. As this practice becomes more commonplace, there will be a need to improve the usefulness of models to demonstrate water quality impacts from stormwater management activities.

To this end, there was an acknowledged need for guidance on how to effectively calibrate stormwater management models. Before relying on models as an alternative to widespread monitoring, there needs to be proper calibration to instill as much confidence as possible. Workshop participants had questions as to how many locations or which activities need to be monitored to provide sufficient data for calibrating a useful model. What is the optimal density of monitoring to inform modeling; is it a cost-effective approach? Workshop participants from southern California indicated that they are transitioning toward relying more on models for predicting water quality impacts because they are responsible for hundreds of water bodies impaired by a wide variety of pollutants. However, they had concerns about the accuracy of water quality outputs from current models.

### 4.4.4 Evaluate Methods to Account for True Source Controls in Models

Participants at both workshops acknowledged a need for better effectiveness data related to source controls and better methods for accounting for such data in stormwater modeling. Since source control is preventative in nature and not treatment-based, it is often difficult to accurately quantify the impact total or partial removal of a specific source has on the quality of a water body. Typical stormwater management models only account for pollutant removal after the occurrence of a rainfall event (e.g., pollutants are already on the ground and are transported via runoff into conveyances and structural BMPs). True source controls remove pollutants from the environment before they have a chance to contact runoff. Several workshop participants expressed the belief that true source control is the most effective BMP and contributes greatly toward meeting regulatory goals like TMDL wasteload allocations. There was an acknowledged need for finding better ways to represent these impacts in predictive models.

## 4.5 Improving Program Tracking and Reporting

**Ninety-two percent** of respondents agreed that *"Reporting requirements should move beyond passive activity and data tallies to incorporate active effectiveness evaluation and clear linkages to program action."*

Tracking and reporting are often discussed in tandem, yet it is important to differentiate between these activities. As part of their NPDES permit requirements, programs must report on their implementation or effectiveness every year. Programs are therefore compelled to perform tracking activities to fulfill this requirement. Since the quality of a tracking program is not evaluated as part of the regulatory obligation, this time- and resource-intensive endeavor can amount to little more than a "bean-counting" exercise if not structured properly. The voluminous paper reporting is another common criticism, especially in programs where NPDES permitting authorities are not able to fully review the annual reports.

Workshop participants indicated that tracking and reporting should have a clear link to the required program activities to enable a true effectiveness assessment. The forthcoming NPDES Electronic Reporting Rule, which requires entities to electronically submit specific permit and compliance monitoring information instead of filing paper reports beginning in 2020, presents a key opportunity to re-envision how tracking and reporting can yield more useful and usable data. It should be noted that 88 percent of survey respondents agreed that e-reporting will not improve reporting quality unless more measurable and evaluative metrics are associated with program activities.

### 4.5.1 Identify an Approach for Using Established Performance Metrics to Guide Tracking and Reporting Efforts

Section 4.1.3 described the need to establish key performance metrics for more effective program monitoring. The Phase II MS4 regulations introduce the concept of establishing "measurable goals" as a component of stormwater management programs to "evaluate the effectiveness of individual control measures and the storm water management program as a whole" (EPA, n.d.). EPA's *Measurable Goals Guidance for Phase II MS4s* explains that there are various ways local programs can write their measurable goals and identifies the following main categories: (1) tracking implementation over time, (2) measuring progress in implementing the BMP, (3) tracking total numbers of BMPs implemented, (4) tracking program/BMP effectiveness, and (5) tracking environmental improvement. Some of these loosely align with the six CASQA outcome levels, with the highest outcome (or measurable goal category) related to improvement in receiving waters. However, measurable goals for most Phase II MS4 programs tend to be more focused on tracking the occurrence of activities or outputs rather than outcomes (categories 1 to 3). The programs then report on a myriad of program activities in their annual reports, which can be cumbersome, time-consuming, and may only provide minimal insight into the effectiveness of the underlying programs.

**Commented [BJ14]:** Note to EPA: Should the Remand Rule for "clear, specific, and measurable" also be mentioned in here? The intent of this can make it easier to see whether program are implementing (and counting the activities) but still is challenging to assess effectiveness from an outcome standpoint.

Dynamic activity tracking, evaluation, and reporting system enables more coordinated program management and adjustment and clearer permit reporting. Focusing on program elements that are linked directly to quantifiable water quality outcomes (e.g., BMP maintenance) and reporting tools that provide transparent accounting of benefits and are field-verifiable will accelerate progress and provide useful information to decision makers. Once a program determines what elements needs to be monitored, it should seek to adopt a more integrated information and data management system that synthesizes data geographically and supports real-time management decision-making. An increasing number of programs are beginning to adopt asset management approaches for integrating

disparate data systems.[5] One workshop participant noted that implementing a more holistic asset management approach provides an appropriate framework for systematic performance tracking. This in turn can promote a better understanding of the correlation between activities to outcomes and generate actionable information on overall performance.

Workshop participants stressed the importance of tracking locations, capacity, types, and performance (or maintenance status) of structural BMPs. Collectively, these serve as useful metrics for determining program progress and permit compliance on short time frames and can guide action prioritization. Another participant noted that tracking and reporting on receiving waters is critical for effective program management and the public engagement. Training and examples will be needed to assist communities in implementing new methods and incorporating them in permits.

### 4.5.2 Determine the Most Effective MS4 Program Reporting Mechanisms and Formats

Improving the functionality of reporting mechanisms will help streamline the process for program staff, making them more likely to fully engage in the effort. Workshop participants suggested that a national stormwater organization (e.g., WEF, NMSA) could survey states to identify the most effective reporting mechanisms currently in place. The results could then be used to inform the development of a Web-based template for implementation under the new E-Reporting Rule. Baseline components would likely include data on receiving waters, outfall monitoring, and interim progress on milestones towards water quality requirements (e.g., wasteload allocation progress for TMDL compliance). Enabling the reporting of more and better data can in turn support the continued development of the local program.

Ultimately, this program information is shared in the annual report. Though the document fulfills a specific regulatory purpose, improving the overall usability would help to promote knowledge transfer across different programs. Workshop participants expressed support for a watershed approach that aggregates information from across the municipalities. Several workshop participants suggested developing a method for an annual report that shows answer and 'work' to benefit multiple audiences. They described a few exemplary local examples that provide online access for regulators and the public alike to dig into program information. This would necessitate a platform or other mechanism for more robust tracking so that annual reports could be more digestible. Indeed, the need to declutter and slim down annual reports to the essential components was a common refrain.

Future reporting systems should be able to incorporate new information as permit requirements, opportunities, and technology shifts over time while providing outputs that clearly communicate program. Guidance and training on new reporting frameworks and how to incorporate them in permits will be needed to advance reporting approaches at the state and local levels.

[5] Asset management is a means to capture information on stormwater asset location, age, type, condition, maintenance history, and cost to help facilitate long-term planning and budgeting, staffing and workflow analyses, enhanced tracking and reporting, proactive maintenance, development of multi-benefit projects, and visual demonstration of progress with identified service levels. The report from the 2017 MS4 workshop included recommendations to (1) build capacity for asset management and (2) incentivize asset management.

# 5 Opportunities and Next Steps

**Commented [AK15]:** Note to EPA: We'll build out this section in a subsequent version after the report content is stable.

# REFERENCES


- Barco, J. Wong, K. Stenstrom, M. (2008). *Automatic calibration of the U.S. EPA SWMM model for a large urban catchment.* Journal of Hydraulic Engineering. (pp. 466-474).
- California State Water Resource Control Board. (2010). *Guidance for assessing the effectiveness of municipal storm water programs and permits.* Retrieved from URL https://www.waterboards.ca.gov/water_issues/programs/stormwater/docs/effctve_assmnt.pdf
- California Stormwater Quality Association. (2015). *A strategic approach to planning for and assessing the effectiveness of stormwater programs.* Retrieved from URL https://www.casqa.org/resources/stormwater-effectiveness-assessment/guidance-document
- Center for Watershed Protection. (2008). *Monitoring to Demonstrate Environmental Results: Guidance to Develop Local Stormwater Monitoring Studies Using Six Example Study Designs.* Retrieved from URL https://dep.wv.gov/WWE/Programs/stormwater/MS4/guidance/Documents/Monitoring_to_demonstrate_environmental_results2008.pdf
- EPA (U.S. Environmental Protection Agency). (2017). *National stormwater calculator – Tool that helps users control runoff to promote the natural movement of water.* Science in action – innovative research for a sustainable future. (EPA Publication No. EPA/600/F-13/095d).
- EPA (2016). *Storm water management model – An application that helps predict the quantity and quality of runoff in future long-term runoff events within urban areas.* Science in action – innovative research for a sustainable future. (EPA Publication No. EPA/600/F-16/246).
- EPA. (2010a). *MS4 permit improvement guide.* (EPA Publication No. 833-R-10-001).
- EPA. (2010b). NDPES Permit Writers' Manual. (EPA Publication No. 833-K-10-001).
- EPA. (2007). *MS4 program evaluation guidance.* (EPA Publication No. 833-R-07-003).
- EPA. (n.d.). Measurable goals guidance for Phase II MS4s. Retrieved from URL https://www.epa.gov/npdes/developing-ms4-resources
- EPA Region 9. (2018). *Evolution of Stormwater Permitting and Program Implementation Approaches.* Retrieved from URL
- EPA Region 9. (2017). *Developing reasonable assurance: A guide to performing model-based analysis to support municipal stormwater program planning*, linked from "NPDES Wastewater & Stormwater Permits" at.epa.gov/region9. US EPA, Pacific Southwest, Region 9. Retrieved from URL: https://www3.epa.gov/region9/water/npdes/stormwater.html
- Fiorino, D. J. (2006). *The new environmental regulation.* Cambridge, MA: MIT Press.
- NPDES (National Pollutant Discharge Elimination System) Municipal Separate Storm Sewer System General Permit Remand Rule, 40 CFR 122 (2016). Volume 81, No. 237.
- NPDES—Regulations for Revision of the Water Pollution Control Program Addressing Storm Water Discharges, 40 CFR 9, 122, 123, & 124 (1999). Volume 64, No. 235.
- Southern California Stormwater Monitoring Coalition. (n.d.). Home. Retrieved from URL http://socalsmc.org/


**Commented [AK16]:** Note to EPA: Is/will this be posted on your site?

# APPENDIX A: WORKSHOP ATTENDEES

| Name | Organization | Location |
|---|---|---|
| Karen Ashby | Larry Walker Associates | Davis, CA |
| Nicole Beck | 2nd Nature | San Francisco, CA |
| Bethany Bezak | D.C. Water | Washington, D.C. |
| Ellen Blake | EPA Region 9 | San Francisco, CA |
| Sean Bothwell | California Coastkeeper Alliance | San Francisco, CA |
| Eugene Bromley | EPA Region 9 | San Francisco, CA |
| Geoff Brosseau | California Stormwater Quality Association | Menlo Park, CA |
| Seth Brown | Water Environment Federation; Storm and Stream | Alexandria, VA |
| Steve Carter | Paradigm H2O | San Diego, CA |
| Matt Fabry | San Mateo County | Redwood City, CA |
| Wes Ganter | PG Environmental | Golden, CO |
| Chad Helme | Tetra Tech | San Diego, CA |
| Bobby Jacobsen | PG Environmental | Golden, CO |
| Peter Kozelka | EPA Region 9 | San Francisco, CA |
| Keith Lichten | San Francisco Bay Regional Water Quality Control Board | Oakland, CA |
| Chris Minton | Larry Walker Associates | Davis, CA |
| Thomas Mumley | San Francisco Bay Regional Water Quality Control Board | Oakland, CA |
| Randy Neprash | National Municipal Stormwater Alliance; Minnesota Cities Stormwater Coalition; Stantec, Inc. | St. Paul, MN |
| Nell Green Nylen | University of California, Berkeley | Berkeley, CA |
| Matt O'Malley | Coastkeeper | San Diego, CA |
| Elizabeth Ottinger | EPA Region 3 | Philadelphia, PA |
| Gayleen Perreira | California State Water Resources Control Board | Sacramento, CA |
| Renee Purdy | Los Angeles Regional Water Quality Control Board | Los Angeles, CA |
| Dominic Rocques | Central Coast Regional Water Quality Control Board | San Luis Obispo, CA |
| Ken Schiff | Southern California Coastal Water Research Project | Costa Mesa, CA |
| Grant Sharp | Orange County | Orange County, CA |
| Dave Smith | EPA Region 9 | San Francisco, CA |
| Chris Sommers | EOA, Inc. | San Francisco, CA |
| Michael Trapp | MDB | XX |
| Suzanne Warner | EPA Region 1 | Boston, MA |
| Richard Watson | RWA Planning | Mission Viejo, CA |

**Commented [BJ17]:** Note to EPA: Updated this as best I can based on info available currently.

FYI 17 of these 31 people attended the first workshop as well

# Appendix B: Workshop Agenda

## Overview

This workshop is the second of two planned meetings that will focus on the evolution of stormwater programs and permitting requirements. The first meeting (in December 2017) addressed minimum control measures, industrial/construction program requirements, and water quality-based control requirements. This second workshop will focus on municipal stormwater program monitoring, tracking, evaluation, and reporting provisions. We will evaluate experiences to date and opportunities to improve in how we:

- ☐ ***Establish Performance Metrics*** *that form the basis of tangible targets and goals for the program and program elements.*
- ☐ ***Monitor stormwate****r, with an eye toward strengthening the linkage between stormwater program actions and our ability to quantify change in stormwater and receiving water quality,*
- ☐ ***Use other evaluation methods*** *(e.g., measuring surrogate measures, activity metrics, BMP implementation, etc.) with, or instead of, water quality measures,*
- ☐ ***Track program implementation*** *and progress in meeting goals (both water quality and other types of program goals), and*
- ☐ ***Report on program progress*** *and accomplishments to stakeholders and permitting authorities.*

As we did in the December meeting, we will focus to a significant degree on how NPDES MS4 permits can be better structured or restructured to encourage/require more useful, cost-effective approaches and reduce or eliminate less effective methods and requirements. Workshop feedback will be synthesized with other existing research to produce a white paper discussing opportunities to strengthen how MS4 permits and implementation programs address monitoring, tracking, evaluation, and reporting.

## Structure

Throughout the workshop, participants will be encouraged to consider whether and how existing MS4 program requirements concerning monitoring, tracking, evaluation, and reporting add value and to identify ways to improve permit and program effectiveness. To enable these discussions, each session will follow the same general structure:

- ☐ **Conversation starter**. A guest speaker will provide a 5-10-minute overview, outlining the current state of monitoring and assessment, summarizing evolution over time, or sharing a brief example case study. In some cases, more than one conversation starter may speak.
- ☐ **Hypothesis review**. As we did for the prior meeting, we will conduct a pre-meeting survey of participants to test a series of hypotheses concerning the effectiveness of current monitoring, tracking, evaluation, and reporting approaches and permit requirements. We will summarize survey responses to help identify the degree of

agreement or disagreement concerning key lessons learned and improvement opportunities.

- ☐ **Discussion.** The facilitator will then lead in-depth group discussions. For each permit element, we will consider 3 basic questions:

  ***1. How effective has these program tools/requirements been in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?***

  ***2. How can implementation of monitoring, tracking, evaluation, and reporting be improved in the future?***

  ***3. How can permits be improved to facilitate desired changes in monitoring, tracking, evaluation, and reporting?***

- ☐ **Findings/Recommendations**. Each session will be focused to solicit important findings and specific actions to strengthen and improve the corresponding MS4 program/permit element. The workshop will conclude with a recap in an effort to identify areas of agreement and disagreement and issues needing further evaluation before adjourning. The work we do at the workshop will inform preparation of a paper that will summarize our work and hopefully help guide future actions to help improve MS4 permits and programs.

## Key Terms

It is imperative that participants understand and attempt to use a common set of terms. Some of these key terms include:

- **Program Assessment** – Using a combination of methods, an analysis of the overall effectiveness of the MS4 program.
- **Monitoring** – Water quality monitoring typically performed at end-of-pipe, in-stream, or in a receiving water.
- **Evaluation** – A determination if the program element, activity, or an individual BMP is meeting stated objectives and performance metrics.
- **Tracking** – Collecting and compiling information on program implementation.
- **Reporting** – Presenting collected information to (1) assist with compliance determinations, (2) demonstrate adherence with Performance metrics, or (3) disseminate information to stakeholders.
- **Activity** – An action taken by a permittee or a regulated entity within the permittees jurisdiction that may provide a water quality benefit.
- **BMP** – A specific structural or non-structural management practice that is known to provide a water quality benefit.
- **Performance Metric** – a qualitative or quantitative measure of an objective or goal.

- **Activity-based** – A measure of output whose benefit to water quality cannot be clearly quantified.
- **BMP Performance-based** – Monitoring results for a particular BMP or set of BMPs; expressed as pollutant concentration, pollutant reduction, or flow reduction.
- **Water Quality-based** – Monitoring results as determined from samples collected at an outfall, in-stream, or within a receiving water.

Other key terms will be identified and defined during the course of the workshop.

## Agenda

### Wednesday, March 21, 2018

| | |
|---|---|
| **9:00-9:30 am** | **Welcome and Overview of Workshop Agenda** |
| | Tom Mumley, San Francisco Bay RWQCB and Wes Ganter, PG Environmental<br>□ Welcome<br>□ Introductions<br>□ Review of Workshop Purpose and Agenda |
| **9:30-10:45 am** | **Session 1: Current Condition - Are the current Monitoring, Evaluation, Tracking and Reporting requirements effective?** |
| **Conversation Starters:** Dave Smith (EPA Region 9) and Grant Sharp (Orange County)<br>*The objective of this retrospective session is to hear positive perspectives on the usefulness of current monitoring, evaluation, tracking and reporting requirements and to identify elements that are working well.*<br><br>**Discussion:** *How effective has these program tools/requirements been in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?* | |
| **10:45-11:00 am** | **Break** |
| | |
| **11:00-2:00 pm** | **Session 2: How Can We Better Use Performance Metrics To Facilitate Improved Monitoring, tracking, evaluation, and reporting?** |
| **Conversation Starters:** Nicole Beck (2nd Nature) and Dominic Roques (Central Coast Regional Water Board)<br><br>**Discussion and Development of Findings and Recommendations**<br>*1. Is it feasible to develop Performance Metrics for the Program and program elements and will this be helpful in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?* | |

| | |
|---|---|
| *2. Does the proposed construct and use of Activity-based, BMP-Performance-based, and Water-quality based Performance Metrics make sense? If not, what other approaches should be considered?*<br><br>*3. How can permits be improved to facilitate desired changes?* | |
| **12:30-1:15 pm** | **Obtain Lunch + Special Attraction- WEF's Stormwater Testing and Evaluation for Products and Practices (STEPP) initiative (Seth Brown, WEF)** |
| | |
| **1:15-2:00 pm** | **Continuation of Session 2 -** |
| | |
| **2:00-2:30 pm** | **Break** |
| | |
| **2:30-4:15 pm** | **Session 3: How Can We Make Outfall and Receiving Water Monitoring More Useful?** |
| **Conversation Starters:** Ken Schiff (SCCWRP) and Chris Minton (Larry Walker & Associates)<br><br>**Discussion and Development of Findings and Recommendations:**<br>*1. How effective has monitoring program tools/requirements been in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?*<br><br>*2. How can implementation of monitoring and evaluation be improved in the future?*<br><br>*3. How can permits be improved to facilitate desired changes in monitoring and evaluation?* | |
| | |
| **4:15-4:45** | **Review of Day 1 and Initial Synthesis** |
| | |

THURSDAY, MARCH 22, 2018

| | |
|---|---|
| **8:30-8:45** | **Reset and Chart Day 2**<br>Wes Ganter, PG Environmental |
| | |
| **8:45-10:00** | **Session 4: Linking Activities To Expected Water Quality Outcomes** |
| **Conversation Starter:** Bethany Bezak (DC Water)<br><br>**Discussion and Development of Findings and Recommendations:**<br>*1. How effective has these program tools/requirements been in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?*<br><br>*2. How can implementation of models and linked planning, monitoring, and data collection methods improve evaluation techniques in the future?*<br><br>*3. How can permits be improved to facilitate desired changes in evaluation?* | |

| | |
|---|---|
| **10:00-10:15 am** | **Break** |
| **10:15-11:30 am** | **Session 5: How Can We Improve Program Performance Tracking?** |
| **Conversation Starter:** Randy Neprash (NMSA)<br><br>**Discussion and Development of Findings and Recommendations:**<br>*1. How effective have tracking tools/requirements been in improving water quality, telling the story about what program effectiveness, and achieving other program objectives?*<br><br>*2. How can implementation of tracking be improved in the future? Are asset management programs the desired solution?*<br><br>*3. How can permits be improved to facilitate desired changes in tracking?* | |
| **11:30-12:30 pm** | **Lunch: Special Attraction: Using Real Time Controls To Optimize Stormwater Management (Chad Helmle, Tetra Tech)** |
| **12:30-1:45 pm** | **Session 6: Reforming Reporting Approaches To Help Move Programs Forward and Give Permitting Authorities What They Need** |
| **Conversation Starter:** Elizabeth Ottinger (EPA Region 3- Philadelphia)<br><br>**Discussion and Development of Findings and Recommendations:**<br>*1. How can implementation of reporting be improved in the future?*<br><br>*2. How can permits be improved to facilitate desired changes in reporting?*<br><br>*3. Is there a model reporting format(s) that can be used as an example or template for programs and permits?* | |
| **1:45-2:15 pm** | **Break** |
| **2:15-4:00 pm** | **Session 7: Reflection, Synthesis, and Wrap Up** |
| ☐ Identify areas of agreement, disagreement, or warranting more exploration.<br>☐ Review and fine tune findings and recommendations.<br>☐ How do we build capacity to use improved methods and approaches?<br>☐ How can we best bring about desirable change in permitting approaches (next steps)? | |
| **4:00-4:30pm** | **Meeting Evaluation and Closing** |

# APPENDIX C: PRE-WORKSHOP QUESTIONNAIRE RESULTS

This paper reports survey results. On the right side of each table, responses were summarized and shaded in cases where responses were particularly strong in one direction or the other, or very balanced. Please keep in mind this is not intended to be a statistically valid survey instrument. Thank you for your responses.

**Key to Shading**

- 80% or more agree or see improvement opportunity
- 70-79% agree or see improvement opportunity
- Even, <20% difference
- 70-79% disagree
- 80% or more disagree

**1. Effectiveness evaluations, program tracking, and reporting - Assuming it is possible to improve and adjust these activities, how would you rate the potential for significant improvement (toward cost-effective environmental outcomes) for each element?**

| | Significant potential | Some potential | Little potential | No potential | No opinion or insufficient knowledge | TOTAL | Significant or Some Potential | Little or No Potential |
|---|---|---|---|---|---|---|---|---|
| Water Quality Monitoring (receiving water, outfall, within collection system, at project or practice scale) | 19 | 5 | 0 | 0 | 0 | 24 | 100% | 0% |
| Non-Water Quality Evaluation (activity evaluation, effectiveness evaluation) | 15 | 6 | 2 | 0 | 1 | 24 | 88% | 8% |
| Tracking (tracking discreet activities (e.g. inspections, street sweeping, BMP installation), active asset management planning and tracking) | 12 | 12 | 0 | 0 | 0 | 24 | 100% | 0% |
| Reporting (annual reporting to permit authorities, reporting to public or elected officials) | 16 | 5 | 3 | 0 | 0 | 24 | 88% | 13% |

**2. What are the key elements of program effectiveness?** *(responses copied directly from survey results; not edited for grammar or spelling)*

- 1) Solid definition of performance metrics

2) Metrics that are linked to meaningful outcomes

3) Suite of activities that directly move those metrics in a measurable way

- We don't really know how our effective our programs are, generally. At the end of the day, we should be measuring impacts on water quality, but that has not been a focus for most programs for both political and financial reasons. Until we start to consistently and comprehensively measure performance, we will have no idea of real progress (or lack of progress).
- Key Elements are:
- Enhanced Awareness
- Behavior Change
- Estimating/Modeling Pollutant Reductions
- MS4 Monitoring
- Receiving Water Monitoring
- Clear articulation of the question wanting to answer, including time, space, and degree of change you're wanting to observe
- Clear and concise permit language that provides flexibility to meet water quality standards while requiring robust monitoring to demonstrate compliance.
- Improvements in water quality (both discharge quality and receiving water quality); reduction in pollutant load discharged (either through stormwater treatment or capture); elimination of non-stormwater discharges; elimination of waterbody impairments (and delisting from CWA section 303(d) list)
- Tracking progress of implementation efforts to improve water quality, including reporting of BMPs laid out in a plan (e.g., EWMP, WQIP, GI Plan).
- Clear and measurable performance metrics and the ability to gauge activities and actions versus those metrics; in the case of MS4 there has to be a tie to water quality improvement and/or protection - this is why we invest the time, money, and effort
- Effectiveness measurements that are:
- primarily outcomes (as opposed to outputs)
- appropriate for the specific BMP
- measured as close as possible in time and space to the result of a BMP
- expressed in a meaningful way (e.g., relative (%) as opposed to absolute)
- as appropriate and possible, expressed in lay terms
- We need clear articulation of program requirements, clear methods for associating actions with expected or observed water responses, and clear accountability expectations to ensure the stormwater agency communicates results clearly to the public and the permitting authority.
- Objective, outcome-based performance metrics. Not just checkboxes of "miles of street swept."

- Engagement and expertise at the MS4 level, adequate funding and authority, good asset management
- Ability to show water quality improvement, behavior change, and an overall understanding of the benefits and challenges associated with urban stormwater
- Spatially-explicit, quantifiable information on pollutant loading-reducing structural BMPs and implementation activities
- Close relationship between measured metrics and expected outcomes
- Receiving water quality improvements are the ultimate goal
- The key element of program effectiveness to me is the ability to establish a relationship between the BMP/action/activity and a reduction in pollutant loads.
- Understanding current level of effort (including common definitions to ensure consistent understanding of those efforts)
  Understanding desired outcomes and meaningful and measurable metrics
- What makes for an effective program?
  Effective programs need continual streams of funding. To obtain funding, program managers need the ability to communicate actions and environmental return both pre- and post-spend in formats easy to understand. Money is spent in specific locations. Spatially-based asset management allows implementation optimization and simplifies tracking and reporting.

**3. Are program assessment requirements outdated and ineffectual?**

| | Strongly Agree | Agree | Neither Agree nor Disagree | Disagree | Strongly Disagree | TOTAL | Strongly Agree or Agree | Strongly Disagree or Disagree |
|---|---|---|---|---|---|---|---|---|
| Permits have been relatively inflexible, resulting in retention of less effective monitoring requirements and difficulty in initiating more effective and innovative approaches. | 9 | 13 | 0 | 2 | 0 | 24 | 92% | 8% |
| Permits have failed to include clearly defined performance metrics that can be fulfilled through coherent monitoring and evaluation approaches. | 9 | 12 | 3 | 0 | 0 | 24 | 88% | 0% |
| Permit monitoring and evaluation requirements have failed to adequately consider program size, complexity, and pollutants of concern. | 8 | 8 | 5 | 2 | 1 | 24 | 67% | 13% |
| Stormwater quality monitoring has been largely ineffective in assisting compliance evaluation, problem targeting, and program improvement. | 12 | 9 | 2 | 1 | 0 | 24 | 88% | 4% |
| The stormwater quality monitoring problems are attributable to lack of experimental designs that have well defined objectives, minimize sampling error and constrain the hydrologic variability in stormwater quality. | 7 | 7 | 8 | 1 | 1 | 24 | 58% | 8% |
| Receiving water monitoring has been only moderately effective for trend analysis and assessing attainment of water quality standards. | 5 | 10 | 7 | 2 | 0 | 24 | 63% | 8% |
| Receiving water problems are attributable to the inherent variability in receiving water quality, lack of expertise and time in evaluating collected data, difficulty of associating changes in receiving water quality to watershed sources, and high monitoring costs. | 10 | 7 | 3 | 1 | 3 | 24 | 71% | 17% |
| Making linkages between BMPs and activities and water quality outcomes has been hampered due to stagnant monitoring designs and a lack of defined performance metrics. | 8 | 10 | 3 | 2 | 1 | 24 | 75% | 13% |
| Monitoring data management and analysis systems have not evolved sufficiently to enable effective evaluation and comparison of monitoring results. | 8 | 8 | 5 | 2 | 1 | 24 | 67% | 13% |
| Tracking and reporting frameworks have not been adequately tied to performance metrics which hamper assessment and reduce cost-effectiveness. | 14 | 6 | 3 | 1 | 0 | 24 | 83% | 4% |
| Tracking and reporting frameworks have yet to acknowledge or endorse asset management systems. | 11 | 11 | 2 | 0 | 0 | 24 | 92% | 0% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Program and effectiveness evaluation should not be limited to permittees. The regulators (state and federal) should produce self-evaluations. These evaluations should include input from the full range of stakeholders (including permittees). The results of these evaluations should be made public for widespread distribution. | 6 | 9 | 7 | 1 | 1 | 24 | 63% | 8% |
| The programs for stormwater research have to change. Identifying, describing, and prioritizing research needs must be an open process that includes the full range of stakeholders (including permittees). The process should clearly define the research needs and publicize corresponding grant opportunities. | 10 | 8 | 5 | 1 | 0 | 24 | 75% | 4% |
| An improved process for technology transfer that translates and distributes research results useful for local implementers is needed. | 13 | 9 | 1 | 0 | 1 | 24 | 92% | 4% |

**4. Should we move toward a mix of Activity-based, BMP Performance-based, and Water Quality-based Performance Metrics, tailored to the local program design?**

| | Strongly Agree | Agree | Neither Agree nor Disagree | Disagree | Strongly Disagree | TOTAL | Strongly Agree or Agree | Strongly Disagree or Disagree |
|---|---|---|---|---|---|---|---|---|
| No one monitoring and evaluation method addresses all the assessment needs; multiple approaches tailored to local circumstances are needed. | 14 | 8 | 1 | 1 | 0 | 24 | 92% | 4% |
| If permitees adopt a consistent performance metric-based accounting system (spatial or otherwise), permits can increase emphasis on performance achievement and reduce emphasis on burdensome record keeping. | 10 | 6 | 6 | 2 | 0 | 24 | 67% | 8% |
| It is recognized that permittees or regulators cannot reliably assess program effectiveness at spatial and time scales relevant to management decision making based solely on measured water quality outcomes. | 11 | 7 | 5 | 0 | 1 | 24 | 75% | 4% |
| Program managers and regulators need to continually review and update management/compliance questions to reflect changes in water quality issues and evolution of program approaches to inform monitoring program adaptation. | 8 | 9 | 5 | 2 | 0 | 24 | 71% | 8% |
| Extensive training and outreach for permit writers, program staff and elected officials will be needed to enable local programs to take this approach. | 12 | 6 | 5 | 1 | 0 | 24 | 75% | 4% |
| Asset management systems provide the ability to define and track a wide array of activity-based metrics. | 10 | 10 | 4 | 0 | 0 | 24 | 83% | 0% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mobile enabled platforms are the most efficient way to facilitate and conduct field assessments and monitoring. | 6 | 9 | 6 | 2 | 0 | 23 | 65% | 9% |
| Metrics should enable evaluation not just of what was done, but also of whether those actions were effective. | 16 | 7 | 1 | 0 | 0 | 24 | 96% | 0% |
| Activity-based metrics should only be developed where BMP performance or water quality is difficult or impossible to measure. | 4 | 5 | 2 | 9 | 4 | 24 | 38% | 54% |
| Where programs have completed comprehensive plans identifying specific BMPs (e.g. through reasonable or other modeling), BMP Performance monitoring should be used to assess effectiveness. | 6 | 14 | 1 | 2 | 1 | 24 | 83% | 13% |
| BMP performance monitoring (water quality and/or volume reduction) should be used when stormwater assets are integrated with hydrologic tools to quantify impacts to receiving waters and cumulative BMP benefits. | 6 | 12 | 4 | 2 | 0 | 24 | 75% | 8% |
| Performance-based monitoring (water quality and volume reduction) can be used when BMPs are deployed in series to measure BMP effectiveness, assess maintenance needs, or to educate community stakeholders on program effectiveness. | 5 | 13 | 6 | 0 | 0 | 24 | 75% | 0% |
| Increased sampling of outfalls and locations within the collection system is needed to accurately target pollutant sources and evaluate BMP effectiveness within time scales of interest to permitting authorities and program managers. | 7 | 8 | 4 | 3 | 2 | 24 | 63% | 21% |
| Small systems may not need to perform water quality monitoring if alternative program evaluation and tracking approaches demonstrate effective BMP implementation and maintenance. | 4 | 6 | 8 | 6 | 0 | 24 | 42% | 25% |
| Performance metrics need to be established in concert with improved monitoring designs and methods (as more fully discussed in Session 3). | 8 | 15 | 1 | 0 | 0 | 24 | 96% | 0% |
| Focusing implementation actions and associated monitoring (and possibly even permits) in smaller watersheds or sewersheds improves capacity to evaluate implementation effectiveness and water quality responses. | 8 | 10 | 6 | 0 | 0 | 24 | 75% | 0% |

## 5. How Can We Make Outfall and Receiving Water Monitoring More Useful?

| | Strongly Agree | Agree | Neither Agree nor Disagree | Disagree | Strongly Disagree | TOTAL | Strongly Agree or Agree | Strongly Disagree or Disagree |
|---|---|---|---|---|---|---|---|---|
| Program managers and regulators need to continually review and update management/compliance questions to reflect changes in water quality issues and evolution of program approaches to inform monitoring program adaptation. | 9 | 11 | 4 | 0 | 0 | 24 | 83% | 0% |
| Water monitoring should continue but based on improved design and methods and tighter connection to performance metrics and program objectives. | 13 | 8 | 2 | 0 | 1 | 24 | 88% | 4% |
| Surrogate measures (e.g., fine sediment, flow) are a viable option for reducing analytical costs and increasing power for identifying spatial patterns and changes over time. | 7 | 9 | 7 | 1 | 0 | 24 | 67% | 4% |
| Instream monitoring requirements should be reduced in order to increase monitoring of outfalls, BMP effectiveness, and/or BMP assessments. | 10 | 6 | 4 | 3 | 1 | 24 | 67% | 17% |
| Water quality change detection will be enhanced with accounting of flow conditions coincident with sampling and guidance for how to use flow data to improve analysis | 10 | 7 | 6 | 1 | 0 | 24 | 71% | 4% |
| Monitoring designs must go beyond just data collection methods to include data management, data analysis, and reporting formats that clearly link data collected with Performance metrics. | 13 | 11 | 0 | 0 | 0 | 24 | 100% | 0% |
| New sampling methods (e.g. automated samplers) and designs can yield more reliable data to help answer management questions and assist real-time project and system management. | 8 | 9 | 6 | 1 | 0 | 24 | 71% | 4% |
| Permit language will need to be modified to authorize use of new methods and designs. | 10 | 7 | 6 | 1 | 0 | 24 | 71% | 4% |
| Training and outreach for permit writers, program staff, and elected officials on new methods and designs are needed to familiarize these groups with their benefits and limitations. | 12 | 10 | 0 | 2 | 0 | 24 | 92% | 8% |